The Interoperable Watersheds Network

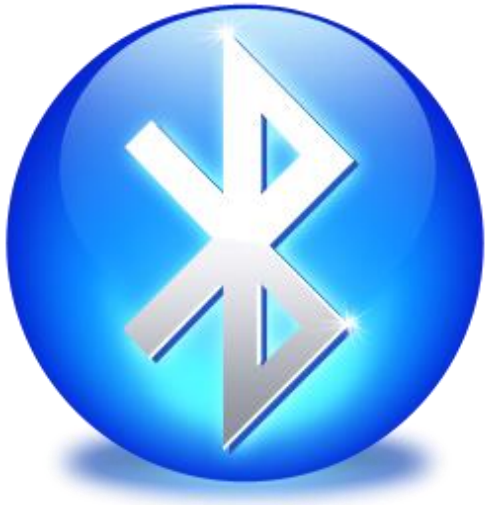
Data Standards in Action

Outline

- What's the problem we're trying to solve?
- What would standardized sensor data enable?
- How IWN is a step towards solving these problems
  - Data Standards
  - Metadata
  - Architecture
- Project Successes/Lessons Learned
- Example: Demonstration Application – CURRENTS
- Example: Mobile App
- Next Steps

The World is Built on Standards!

- When people agree to do things in a common way it opens up many opportunities for everybody to do things better.
  - Technology standards: like Bluetooth and Wifi
  - Data Standards: like financial exchanges
  - Government Standards: like the Exchange Network



Environmental Information eXchange Network

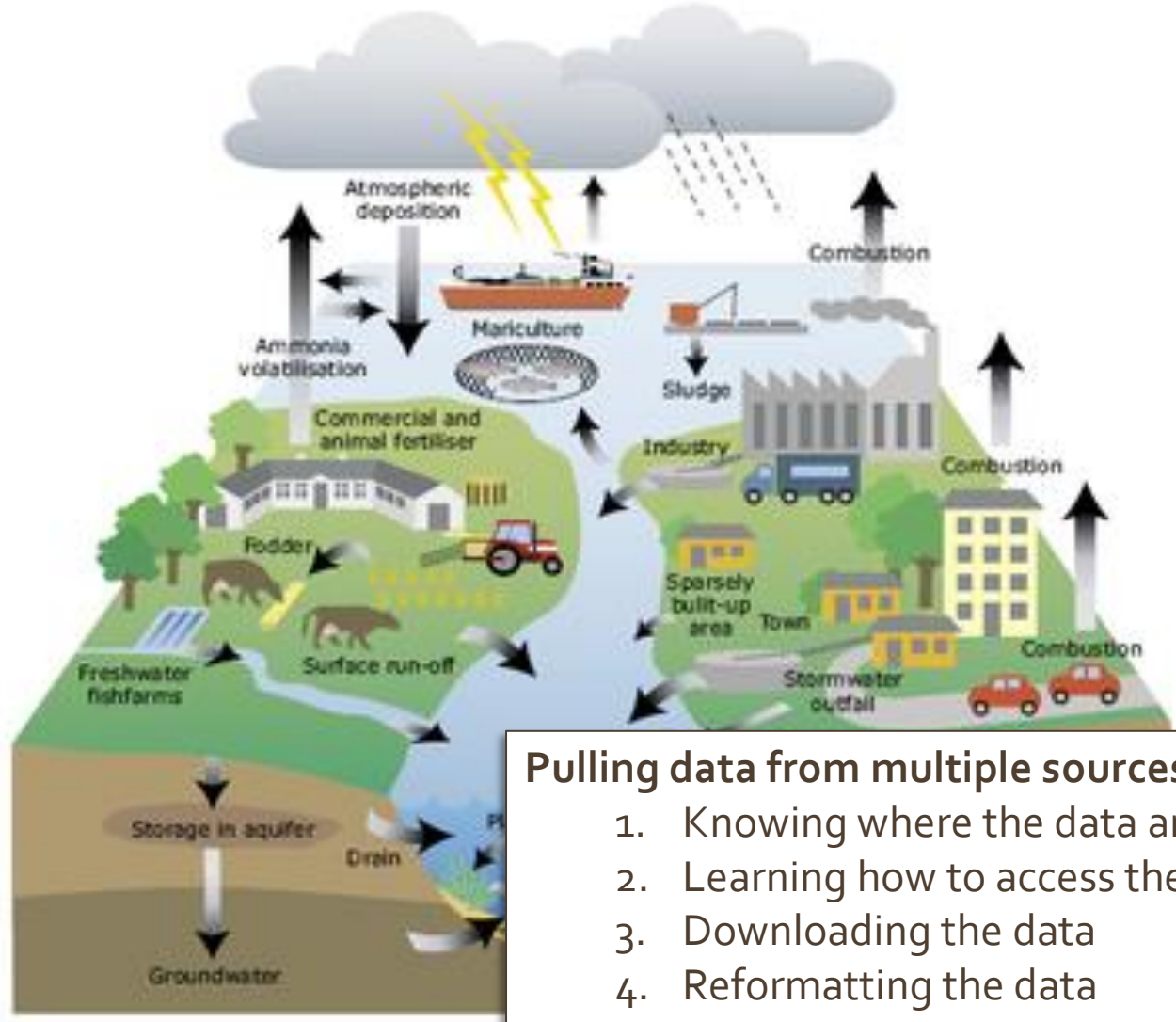
Other Examples of Standards

The IRS has open standards for electronically filing your taxes, enabling a better user experience

Banking: OFX Standards enable you to access your banking information from many 3rd-party applications

Why Do We Need a Sensor Data Sharing Network?

- Water sensors are emerging as a key technology that can be used to improve monitoring efforts
- Multiple entities (EPA, other federal agencies, states, tribes, local groups) are investing in these new technologies
- This has already resulted in a proliferation of data that are not interoperable



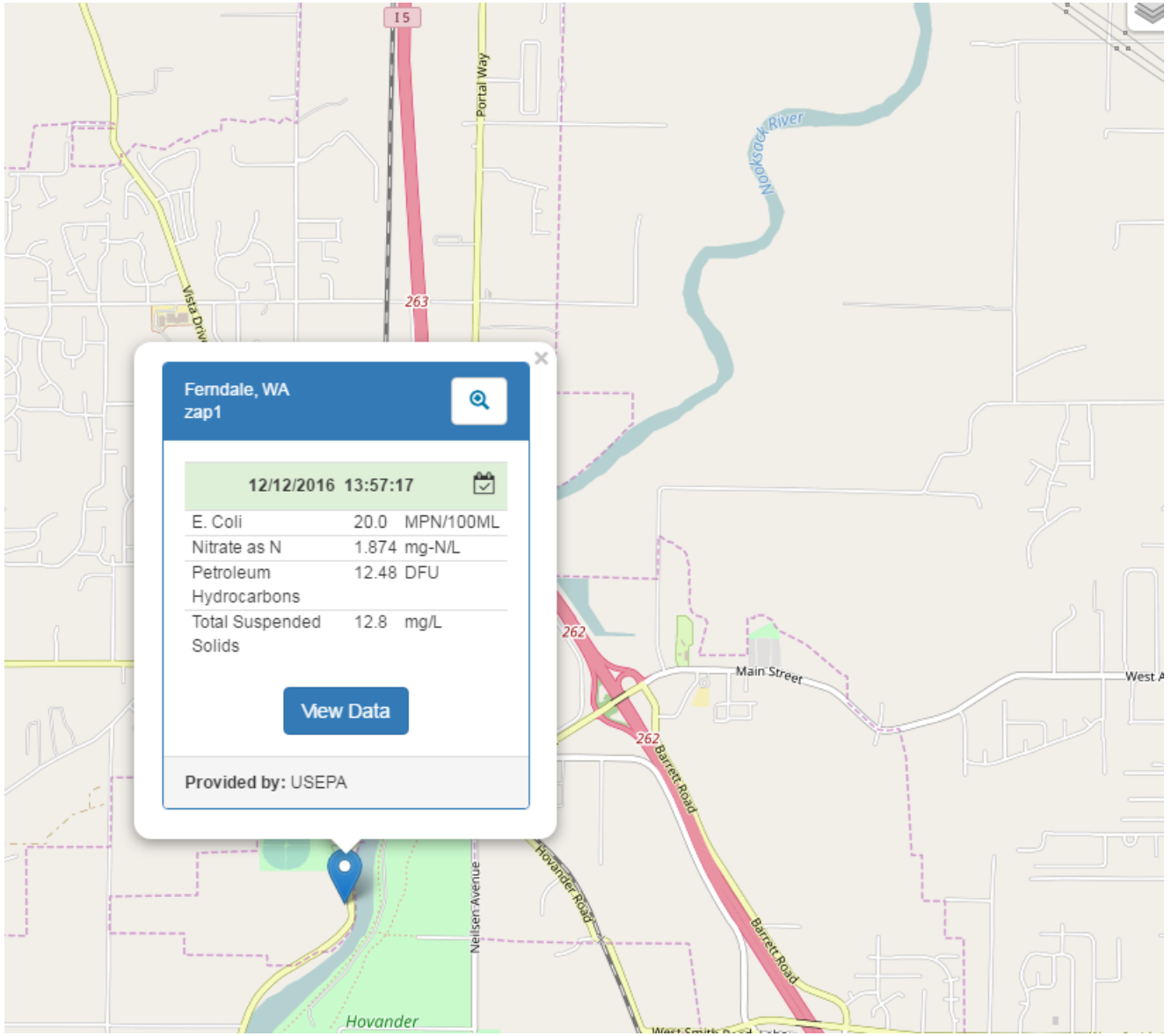
Pulling data from multiple sources requires:

1. Knowing where the data are stored
2. Learning how to access the data
3. Downloading the data
4. Reformatting the data
5. Harmonizing Terms

This results in a significant amount of lost time

How IWN is a Step Towards Solving These Problems

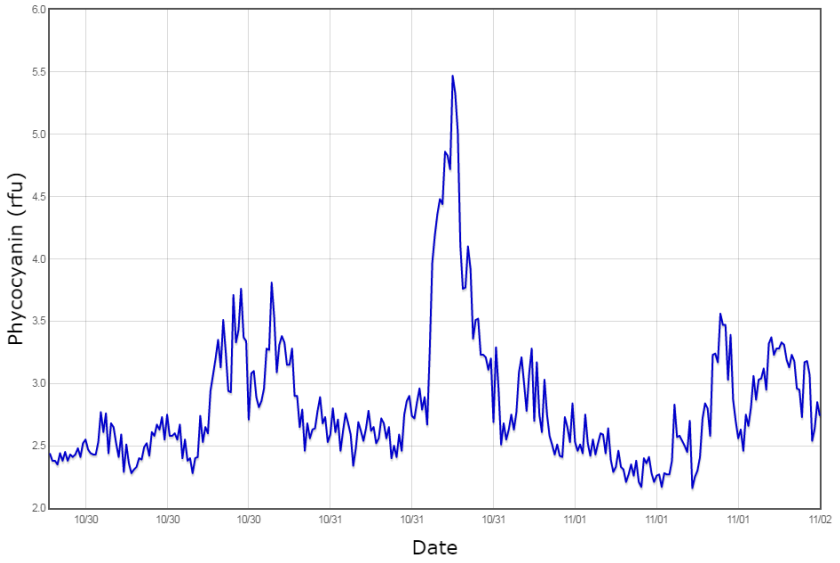
- The Interoperable Watersheds Network was a demonstration project that focused on evaluating approaches to improve sensor data sharing
- It was based on knowledge gained from a recommendations report that EPA developed in 2014
- The project focused on addressing three major areas:
  - Data Standards
  - Metadata
  - System Architecture



The Data Standards Problem

- We needed a common way to represent and communicate the data
- Standards for sensor data already exist, there was no need to create new standards
  - OGC Sensor Observation Service
  - OGC Water ML 2 and Sensor ML
- The Open Geospatial Consortium is an open-source, international standards setting body

OGC®
Making location count.



The Metadata Problem

- Needed a standard way to answer the following questions:
  - What data are available and for what parameters?
  - What data can I use?
  - What's the quality of the data?
- IWN had to develop standard ways to do this (no existing standard existed)
- Further work needs to be done in this area

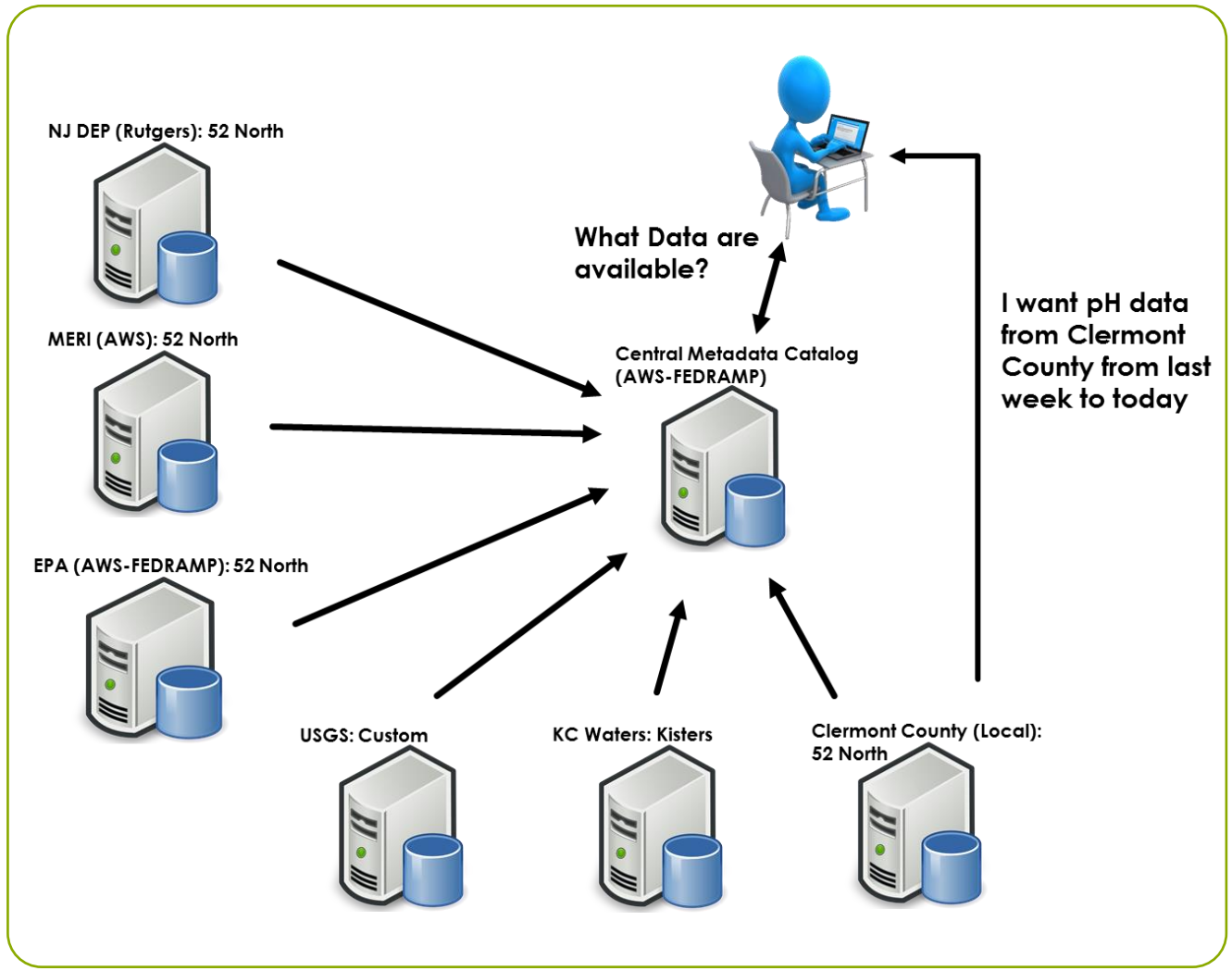
▶ zap1 ✖

- nitrate* (11/10/2016 - 02/13/2017)
- oil* (11/10/2016 - 02/13/2017)
- total_suspended_solids* (11/10/2016 - 02/13/2017)
- e_coli* (11/10/2016 - 02/13/2017)

IWN Used a New Approach for Sharing Data

How do you solve the problem of multiple data providers with large amounts of data that have the potential to change every 3-15 minutes?

- Used a central catalog/index that references every data owner's assets with the corresponding metadata for each sensor
- Allowed for quick searching and discovery of available data
- This approach is similar to how Google allows you to search the internet
- Actual data comes from the partners systems in real-time



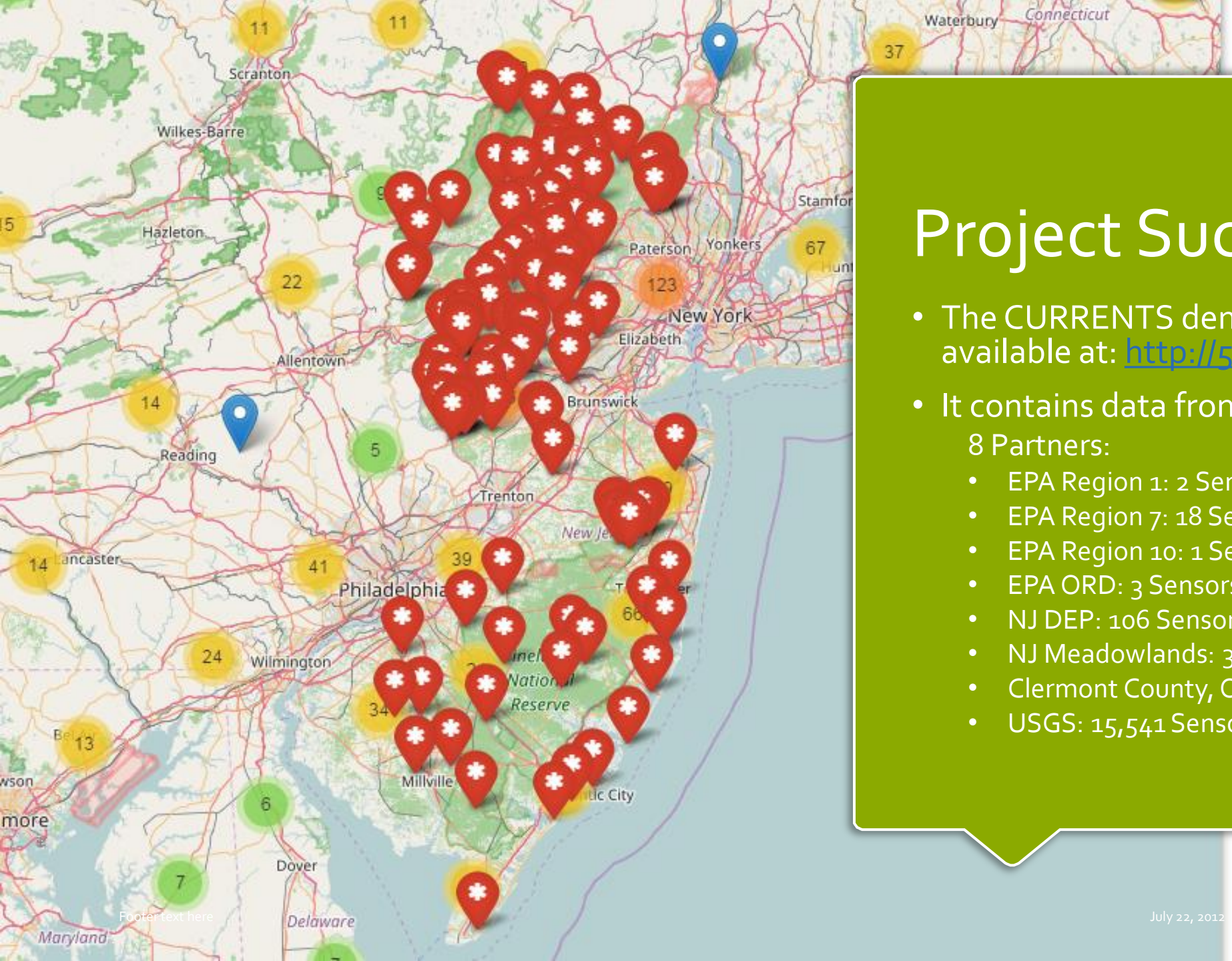
Underlying Catalog Services IWN Defined

- **GetOrganizations**: Returns who is providing data with their endpoints
- **AvailableParameters**: What parameters can be queried
- **GetSensors:** Gets the sensor information and provides different methods for querying sensors (i.e. by county, by HUC, by buffer, by a bounding box, upstream, and downstream)
- **GetSensorParameters**: Gets parameters for a sensor, including the period of record
- **GetOrganizationParameters**: Gets the parameters for an entire Organization

OGC Defined Services (SOS 2)

Each endpoint supports Sensor Observation Service in XML format (WaterML2)

- **GetObservation**: Gets the most recent data or retrieves a collection of data
- **GetCapabilities**: Getting all the metadata from the endpoint
- **DescribeSensor**: Describes the sensor



Project Successes

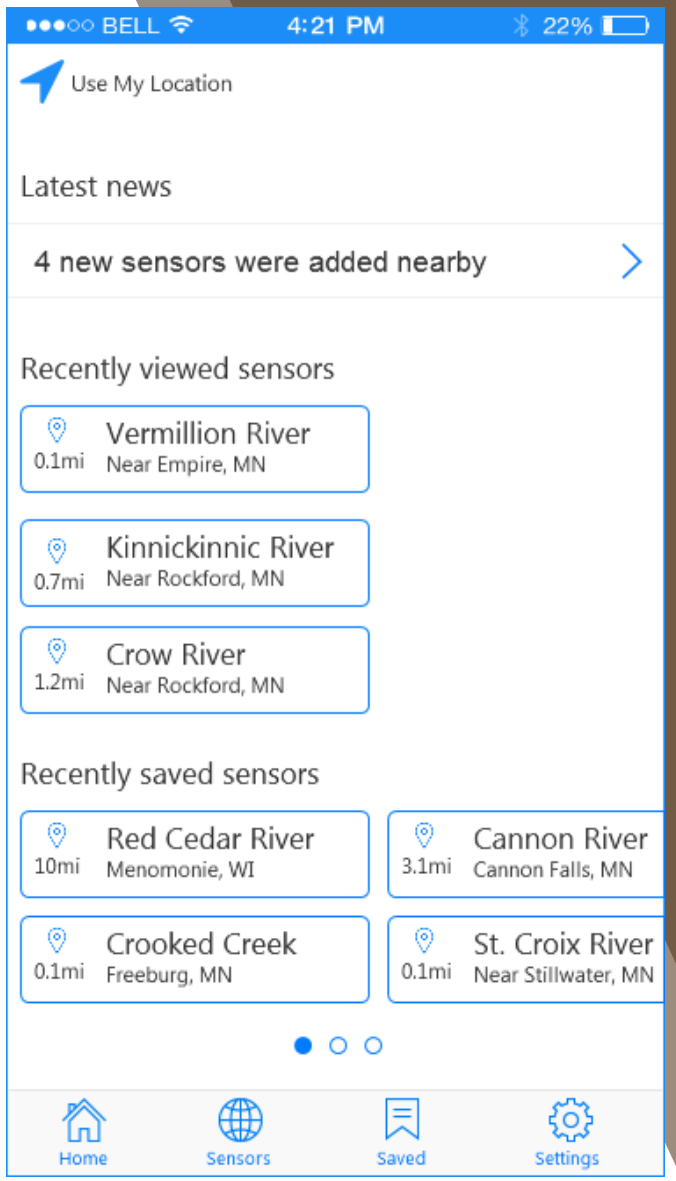
- The CURRENTS demonstration tool is available at: http://54.210.62.171
- It contains data from:

  8 Partners:
  - EPA Region 1: 2 Sensors
  - EPA Region 7: 18 Sensors
  - EPA Region 10: 1 Sensor
  - EPA ORD: 3 Sensors
  - NJ DEP: 106 Sensors
  - NJ Meadowlands: 3 Sensors
  - Clermont County, OH: 4 Sensors
  - USGS: 15,541 Sensors (nationwide coverage)

Organizations

an0605

an0742

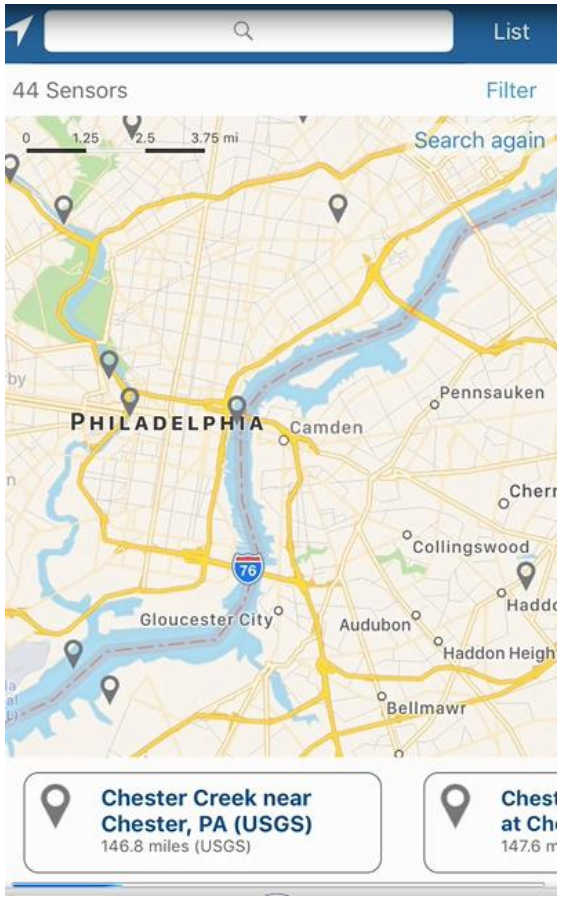


IWN's Open Architecture Allows Other Possibilities

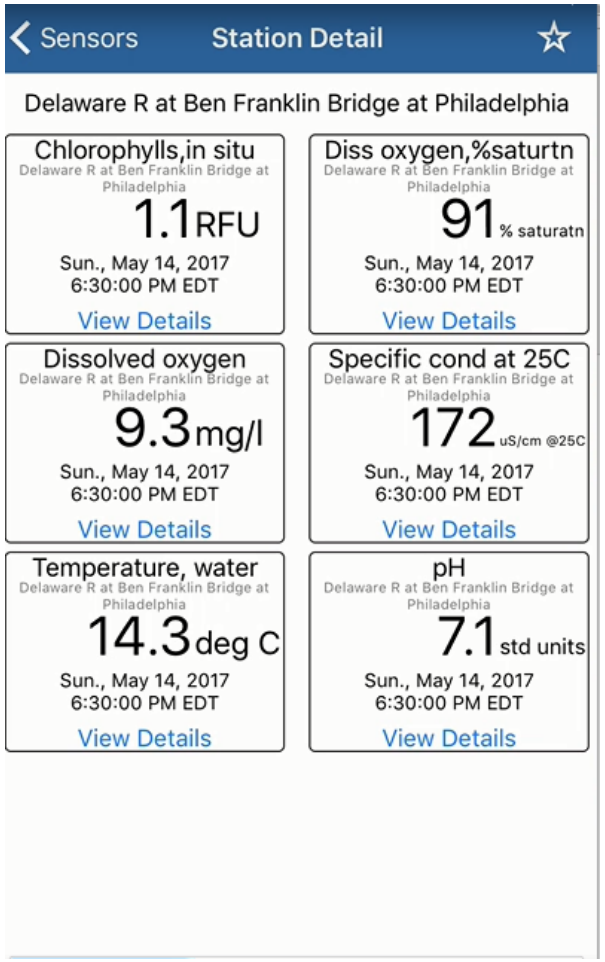
- IWN is built using an open architecture, meaning that all the functionality you see in the demonstration tool is also available as a corresponding **Web Service** or **Application Program Interface (API)**
- Enables other apps to be developed (like mobile apps)
- Also allows other third-party applications (like Excel) to be able to directly interact with the data without having to go to a website and 'download' the data

Real.m in Action

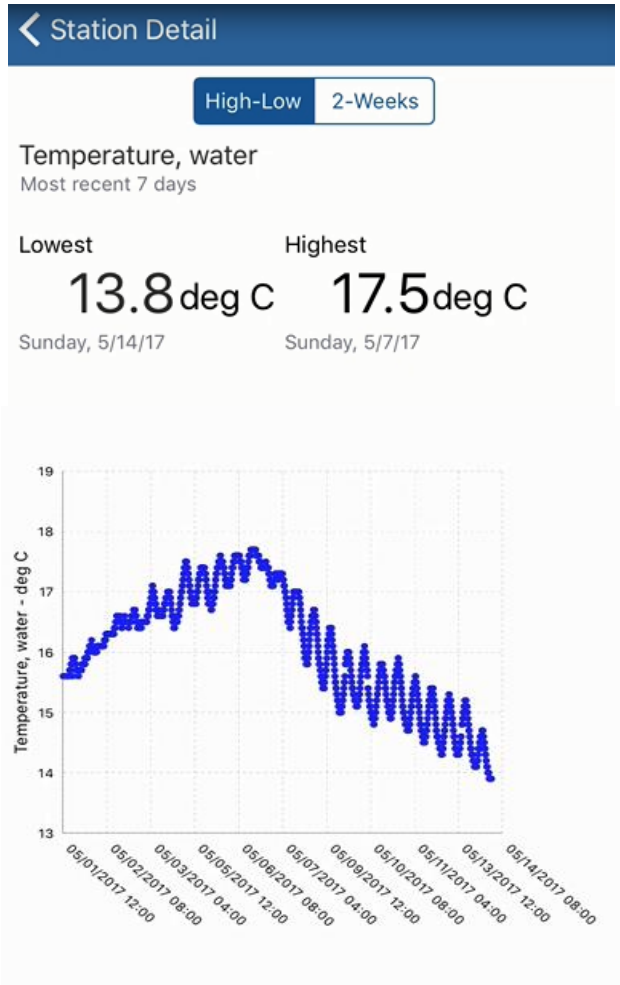
Discover Sensors



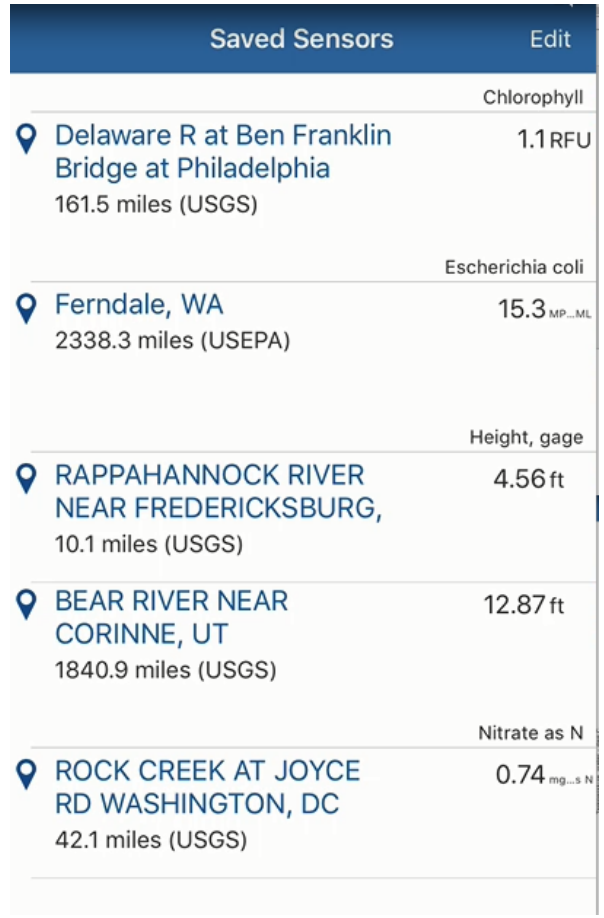
See Current Readings



Understand Trends



View Favorite Sites



Next Steps

- Demonstration project ended in December 2016
  - A Lessons Learned Report has been completed
    - Available at: https://www.epa.gov/sites/production/files/2017-01/documents/iwn_lessonslearned_final_201612.pdf
  - Demonstration tool will continue to be available
  - A mobile app was developed that leverages the services/API developed as part of this project
- Demonstration proved successful
  - Services worked better than expected
  - Setting up a data appliance was simpler than anticipated
- Working with Region 10 to improve the visualization and to make it easier to register sensors and flow data

Improved Sensor Registration Process and QAQC implementation

Progress to date:

- Three Sensors are currently flowing data to Currents:
  - -Nooksack at Ferndale; Nooksack at Lynden; and Fishtrap and Lynden
- Data Appliance complete:
  - Data ingestion software complete and ready for deployment
  - Allows for optional QAQC implementation
  - Available on EPA GitHub: https://github.com/USEPA/Interoperable-Watersheds-Network-Data-Appliance
- QAQC implementation:
  - Tests for values outside of range and removes values outside of range while also keeping the raw data

Future Tasks:

- Deploy to OEI Amazon Cloud

Interoperable Watersheds Network

Organization

Organization Identifier:	epa
Organization Name:	U.S. EPA
Organization Subpart Identifier:	epar10
Organization URL:	www.epa.gov
Contact Name:	Anne Dalrymple
Contact Email:	dalrymple.anne@epa.gov

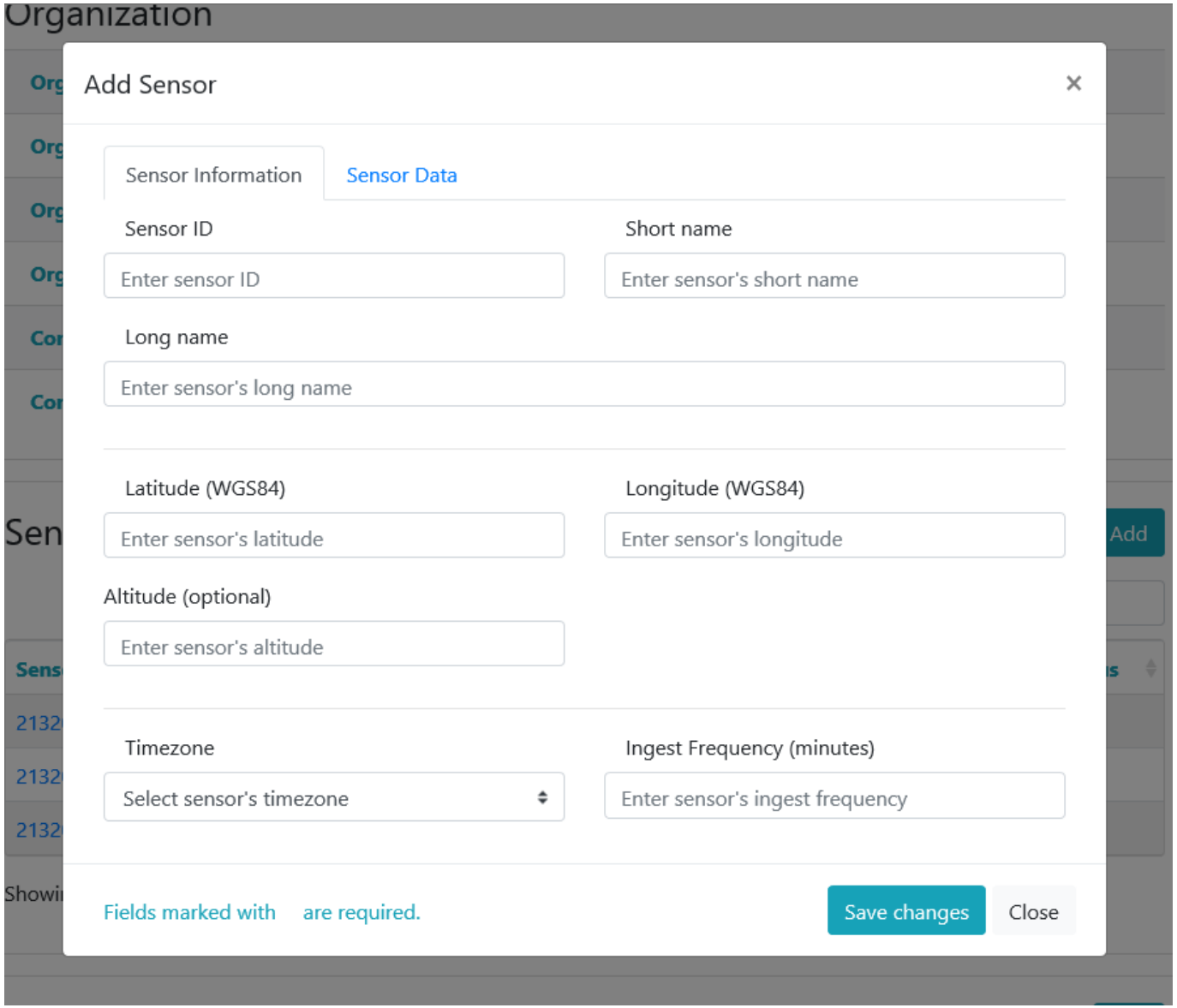
Sensors

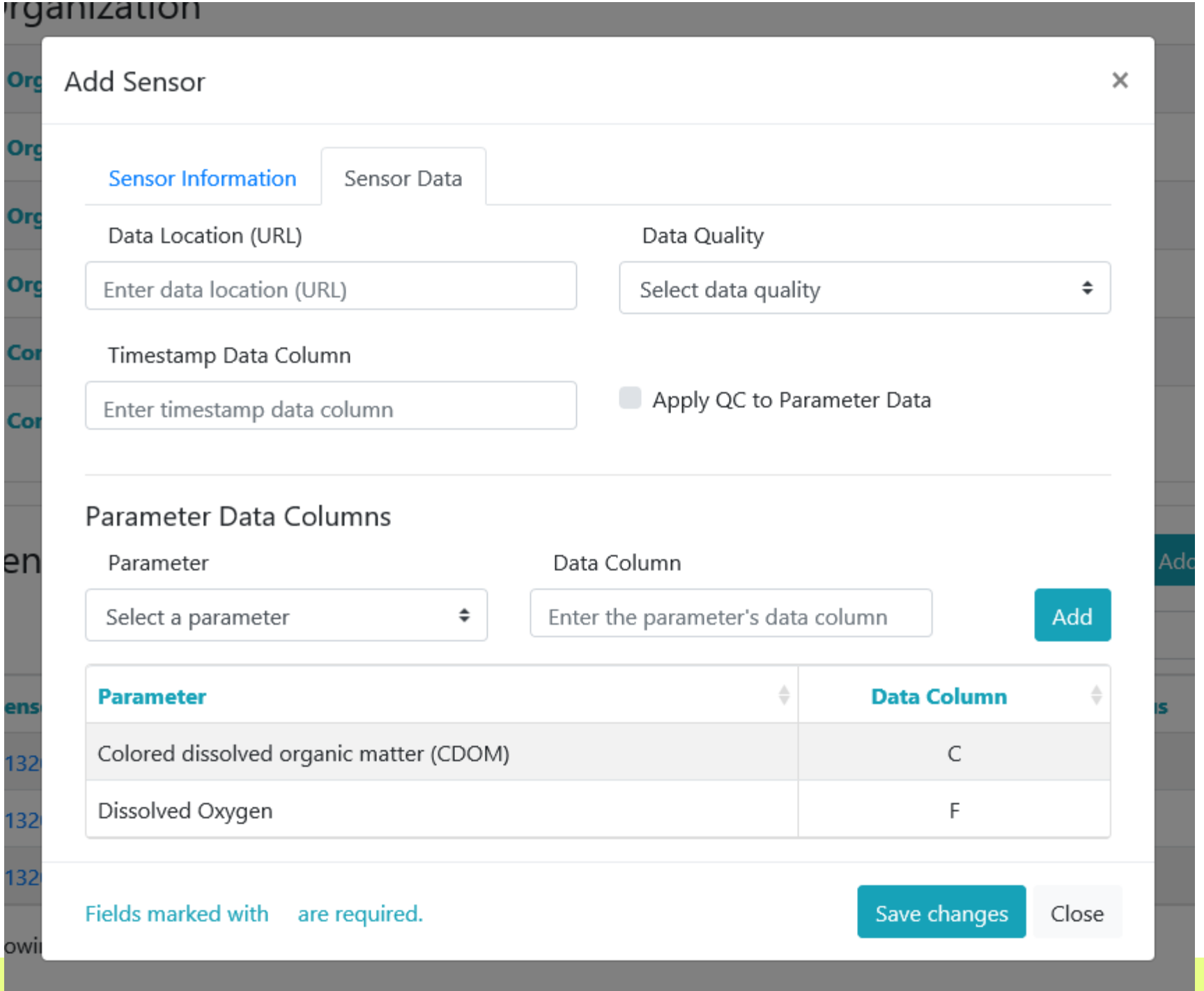
Add

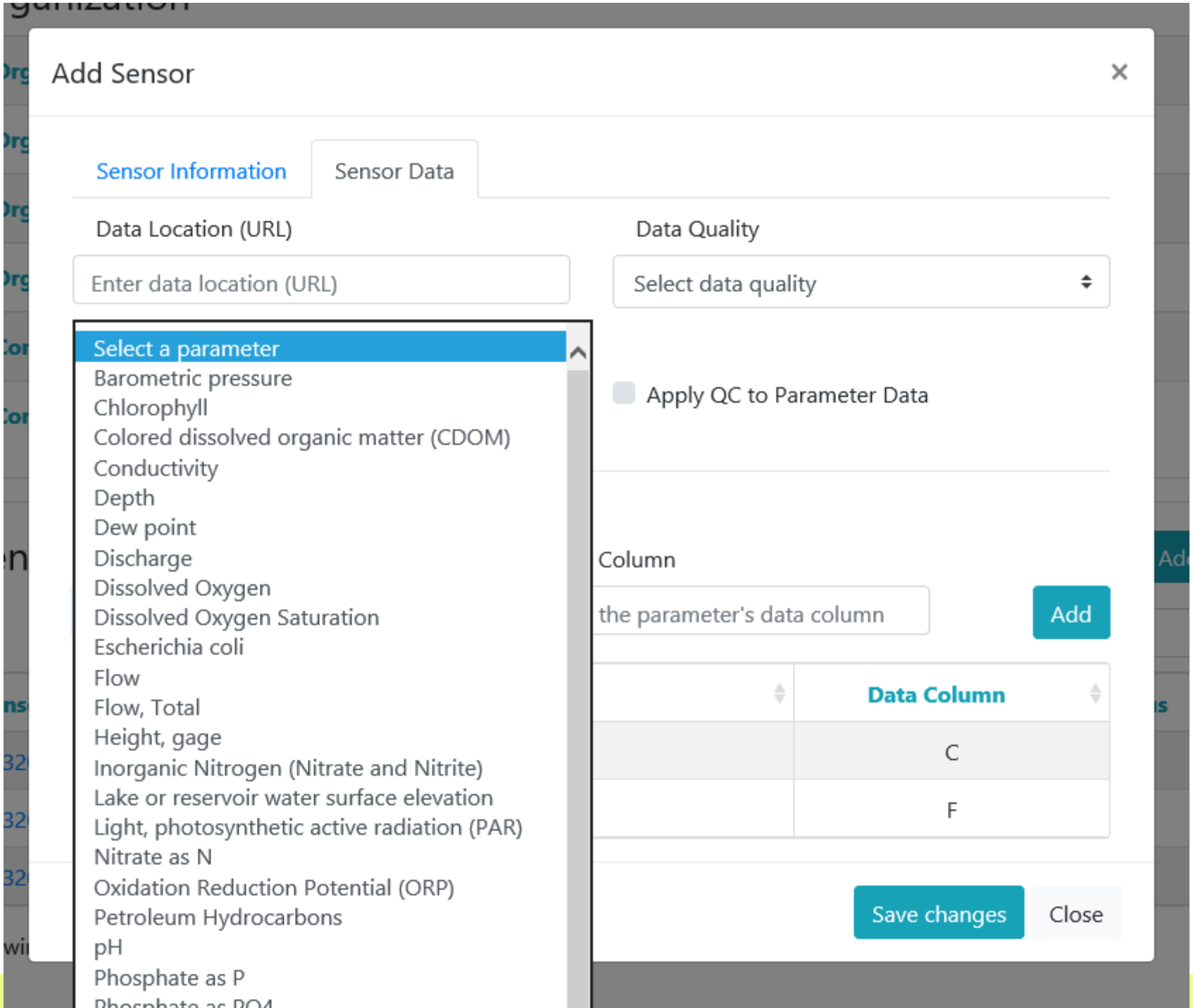
Search

Sensor Id	Sensor Name	Ingest Frequency	Last Ingested	QC	Status
213204	Nooksack@Ferndale	15 min	2018-04-13 04:00:41.975126		
213205	Nooksack@Lynden	15 min	2018-04-13 04:00:41.975126		
213206	Fishtrap@Lynden	15 min	2018-04-13 04:00:41.975126		

Showing 1 to 3 of 3 rows


Organization
Add Sensor
Sensor Information
Sensor Data
Sensor ID
Enter sensor ID
Short name
Enter sensor's short name
Long name
Enter sensor's long name
Latitude (WGS84)
Enter sensor's latitude
Longitude (WGS84)
Enter sensor's longitude
Altitude (optional)
Enter sensor's altitude
Timezone
Select sensor's timezone
Ingest Frequency (minutes)
Enter sensor's ingest frequency
Fields marked with are required.
Save changes
Close
Add


Add Sensor
Sensor Information
Sensor Data
Data Location (URL)
Enter data location (URL)
Data Quality
Select data quality
Timestamp Data Column
Enter timestamp data column
Apply QC to Parameter Data
Parameter Data Columns
Parameter
Select a parameter
Data Column
Enter the parameter's data column
Add
Parameter
Data Column
Colored dissolved organic matter (CDOM)
C
Dissolved Oxygen
F
Fields marked with are required.
Save changes
Close


Add Sensor
Sensor Information
Sensor Data
Data Location (URL)
Enter data location (URL)
Data Quality
Select data quality
Select a parameter
Barometric pressure
Chlorophyll
Colored dissolved organic matter (CDOM)
Conductivity
Depth
Dew point
Discharge
Dissolved Oxygen
Dissolved Oxygen Saturation
Escherichia coli
Flow
Flow, Total
Height, gage
Inorganic Nitrogen (Nitrate and Nitrite)
Lake or reservoir water surface elevation
Light, photosynthetic active radiation (PAR)
Nitrate as N
Oxidation Reduction Potential (ORP)
Petroleum Hydrocarbons
pH
Phosphate as P
Apply QC to Parameter Data
Column
the parameter's data column
Add
Data Column
C
F
Save changes
Close

Contact Name:

Contact Email:

Sensors

Add

Sensor Id	Sensor N	QC	Status
213204	Nooksack		
213205	Nooksack		
213206	Fishtrap@		

Showing 1 to 3 of 3 rows

Quality Control

Add

Search

Parameter	Operand	Threshold	Action
Escherichia coli	>	5000	Discard

Showing 1 to 1 of 1 rows

Quality Control

Parameter

Dissolved Oxygen

Operand

>

Threshold

200

Action

Select action

Fields marked with are required.

Save changes Close

QUESTIONS?

Dwane Young
Young.dwane@epa.gov
202-566-1214